Advantages and Limitations of Simulating Percussion Gestures for Sound Synthesis

Alexandre Bouënard, Marcelo M. Wanderley, Sylvie Gibet

ADVANTAGES AND LIMITATIONS OF SIMULATING PERCUSSION GESTURES FOR SOUND SYNTHESIS

Alexandre Bouënard *†   *Marcelo M. Wanderley* *   *Sylvie Gibet* †‡

* IDMIL / CIRMMT
McGill University
Montreal, Qc., Canada

† SAMSARA / VALORIA
Université de Bretagne Sud
Vannes, France

‡ Bunraku / IRISA
Université de Rennes I
Rennes, France

ABSTRACT

The increasing availability of software for creating real-time simulations of musical instrument sounds allows for the design of new visual and sounding media. Nevertheless, from a conceptual and pratical point of view, the question of how these new instruments can be controlled has rarely been addressed in the literature. In this paper, we present a framework for the control of virtual percussion instruments by modeling and simulating virtual percussionists, based on a motion capture database and on a physically-based movement simulation environment. We show the benefits and limits of such an approach as a means of synthesizing new expressive percussion performances.

1. INTRODUCTION AND MOTIVATION

During the past decades, the design of new digital musical instruments and sound synthesis methods have been widely studied [12, 21]. Especially regarding percussion-related systems, an important research direction is the development of devices to track performer gestures for controlling sound synthesis processes, such as the Radio Baton [5], the Korg Wavedrum [23] or the ETabla controller [19]. Despite the availability of various devices, the most accurate hardware for tracking percussive gestures remains camera tracking systems [25]. These systems offer an effective method for capturing, analysing and virtually reconstructing the performer's whole body, but they fall short in retrieving the dynamic aspects of playing techniques. Moreover, mapping the recorded motion to sound synthesis processes many times relies on non-intuitive multi-dimensional correspondances [14]. Finally, with such methods it is also far from straightforward to go beyond the recorded data and reuse it to synthesize adaptative and realistic new performances.

At the same time, developments in sound synthesis have given rise to various methods to generate percussive sounds. Specifically, physics-based synthesis of percussive sounds has involved the modeling of a hammer [3], collisions and sliding excitations [4], and drum skins [11]. However, their main limitation seems to lie in the way they are controlled. Despite a few early attempts to approach this issue, it is still not clear how to formally relate these models to the excitation by a (real or virtual) performer instrumental gesture, even with the availability of relevant works regarding the study of instrumental percussion gestures [13] and the design of new percussion controllers [2]. Only a few previous works have explored the modeling of the equivalent gestural actions thanks to a slowly evolving mechanical model [16], and then the simulation of an articulated arm hitting a vibrating membrane [17]. More recent attempts to overcome this limitation include works involving the animation of virtual instrumentalists (or virtual models acting as instrumentalists) [18, 20], and the synthesis of sounds from rigid body simulations [26, 22].

In this paper, we present a framework that combines the use of a motion capture database of real percussion performances with the introduction of a physics-based movement modeling and control method of a virtual percussionist. Our approach differs from the methods for animating virtual characters described above by the presence of a physics layer for animating adaptative and responsive virtual percussionists that control the sound synthesis process. It differs from the contributions aiming at synthesizing sounds from rigid body simulations by focusing on the simulation of instrumental gestures, putting the stress on the mapping between new simulated instrumental gestures and sound synthesis algorithms. This approach allows for the synthesis of new virtual percussion performances.

The paper is organized as follows. Section 2 presents the global architecture of our framework. Details on the simulation of percussion gestures and the mapping to sound synthesis processes are given in section 3. Section 4 presents experiments that have been conducted with our framework, namely the synthesis of percussion gesture units that can be composed for creating new percussion sequences. Finally, section 5 discusses the advantages and limitations of our approach, and concludes with further perspectives.

2. GLOBAL SYSTEM ARCHITECTURE

The proposed system integrates both visual and sound feedback through the real-time simulation of instrumental percussion gestures from a user-defined sequence of gestures, a ”gesture score”.

The global architecture involves four steps: a) the planning of a sequence of instrumental gestures using data from a motion capture database, b) the physics simulation of the chosen instrumental gestures, c) mapping strategies between simulated gesture and sound response, and d) the sound synthesis process (Figure 1).

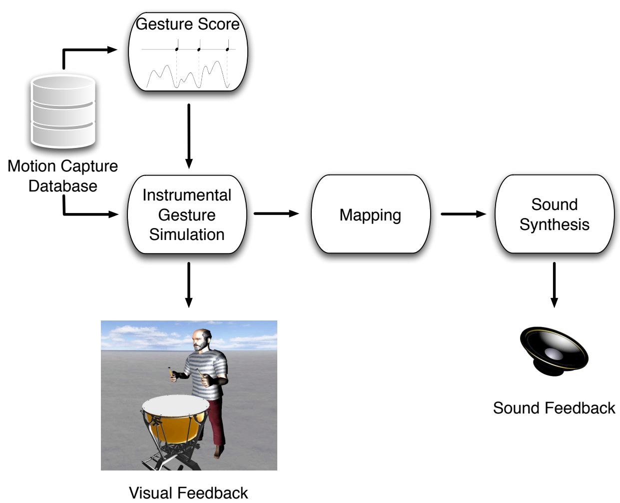


Figure 1. Global system framework.

The planning step takes advantage of a motion capture database of pre-recorded timpani gestures [8]. This database contains the full-body continuous movements of three percussionists performing various exercises following the chosen timpani instrumental gesture typology: different drumstick grips (French or German), beat locations, as well as musical variations (legato, tenuto, accent, vertical accent and staccato).

A refinement to the original motion capture database was implemented by segmenting each multi-beat instrumental performance captured into single-beat units [9]. As proposed in [10], gestural segments are obtained by the physical activity of the interaction instrumentalist-instrument (in our case beat impacts during percussion performances). On the contrary to the generic signal-based framework presented in [10], we rather represent gestural segments using the timpani intrumental gesture typology described above. It allows for the creation of a ”score” at the symbolic gesture level by concatenating several gesture units (Figure 2). Gesture units are then directly derived in signals during the synthesis process.

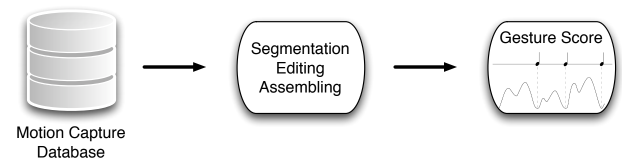


Figure 2. Instrumental gesture planning from the motion capture database: editing a score at the gesture level and assembling gesture units.

The simulation step translates the gesture score into motion trajectory inputs that controls the motion of a biomechanical model of a virtual percussionist (section 3). It is divided in two parts: a) a physics-based environment that handles the simulation of motion equations, as well as the modeling of the virtual percussionist, while b) the control loop drives the motion of the virtual percussionist using information provided by the gesture score.

A mapping layer (section 3.3) then makes a correspondence between outputs from the simulation and inputs of sound synthesis processes. The system currently supports a choice of sound generation options using the Open Sound Control (OSC) protocol, from pre-recorded timpani sounds to physical modeling of membranes.

3. INSTRUMENTAL GESTURE MODELING AND SYNTHESIS

3.1. Simulation and Modeling

The motion of a solid of mass m is described and simulated by solving equations (1) and (2) below at each time step during the simulation. The dynamic forces F_M and torques τ_M applied on a point M of the solid are processed depending on its current state, including its linear acceleration a_M, inertia matrix (or tensor) I_M, and its angular velocity Ω.

$$F_M = m.a_M \tag{1}$$

$$\tau_M = I_M.\dot{\Omega} + \Omega.I_M.\Omega \tag{2}$$

The physics-based modeling of the virtual percussionist is composed of a set of rigid solids linked by mechanical articulations. The biomechanical properties B of these solids (length, mass, density, inertia) are initialized by motion capture data (Figure 3).

3.2. From a score to the physics control of the virtual percussionist

The physics control of the virtual percussionist is achieved by modeling mechanical articulations as damped springs between any two solids, making possible to specify the desired angular state (target angle T) to be reached between two linked solids. These articulations are parameterized by damping and stiffness coefficients (k_s, k_d) that are manually tuned, since there is currently no automatic method for determining them.

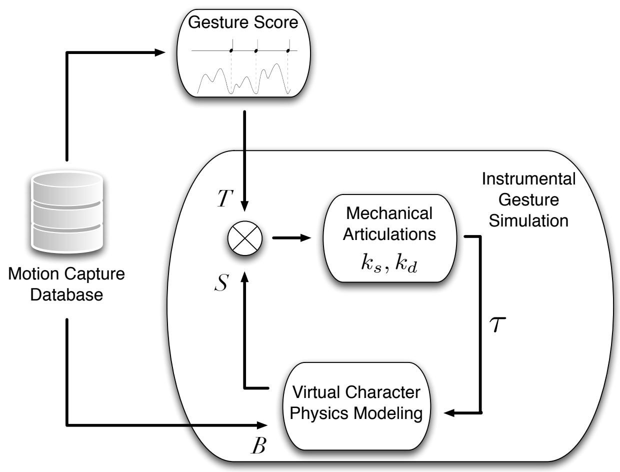


Figure 3. Instrumental gesture simulation: physics modeling and control of the virtual percussionist which aims to reproduce motion trajectories derived from the score.

The physics control loop (Figure 3) involves, for each articulation, its angular state S and angular target T, and processes the torques τ to be applied on the linked solids for reaching the target angle T, as described by equation (3).

$$\tau = k_s.(S-T) - k_d.\dot{S} \tag{3}$$

These target angles T are directly obtained from the gesture score, which is derived into angular motion trajectories at the articulation level.

The presented physics framework for simulating percussion gestures from a gesture score is based on the physics library *Open Dynamics Engine* [24]. The control algorithms we developped and that are at stake during the simulation are presented in [6, 7].

3.3. Mapping between Instrumental Gesture Simulation and Sound Synthesis

The physics layer is exploited to express the mapping between the simulation of percussion gestures and sound synthesis methods at the physics level (Figure 4).

From the simulation of percussion gestures, a collision detection module can retrieve the physics features of any contact event. Considering the interaction between simulated percussion gestures and a physics representation of a drum membrane, the collision detection algorithm provides information on the impact position, velocity and force (direction and amplitude). These features are usually the inputs to physics-based sound synthesis algorithms.

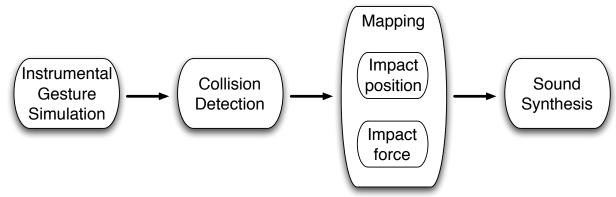


Figure 4. Mapping between instrumental gesture simulation and soud synthesis.

4. RESULTS

We have designed two applications for testing our framework. The first consists of a PureData interface allowing users to choose and explore the simulation of various percussion gesture units and several types of sound feedback. The focus of the second application is the editing and simulation of a score composed of percussion gesture units.

4.1. Simulation of Instrumental Percussion Gesture Units and Sound Feedback

A test-bed application has been designed for exploring the simulation of percussion gestures and sound synthesis processes, with the integration of both visual and sound feedback. Users can parameterize the percussion gesture to be simulated from a Pure Data user interface (Figure 5), choosing among five playing variations (legato, tenuto, accent, vertical accent and staccato), two grip modes or varied tempi. As for visual feedback (Figure 9), users can explore the percussion performance space and visualize some features of interest, such as drumstick trajectories and the resulting beat impact locations.

The user interface also offers various types of sound feedback: pre-recorded percussive sounds, signal-based and physics-based sound synthesis. For instance, the physics-based sound synthesis currently uses a modal synthesis algorithm of a drum membrane from Modalys [1, 15]. Users can parameterize the properties of the drum membrane (radius size, mass and tension) and tune other parameters related to modal synthesis (particularly the number of modes and the resonance). When using this option of sound synthesis, the percussion gesture simulation provides Modalys with the beat impact position and force according to a one-to-one mapping.

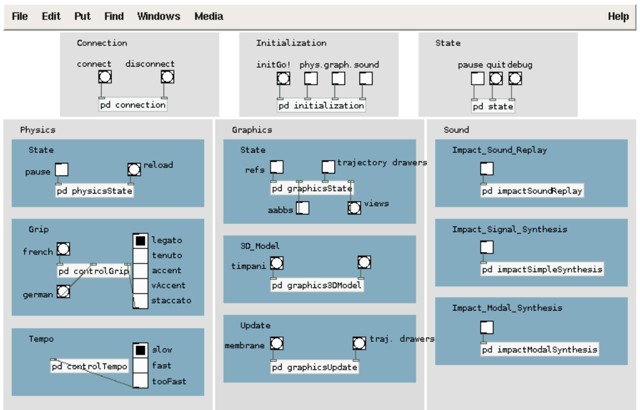


Figure 5. Pure Data user interface: on the left panel users select a drumsctick grip and the playing variation to be simulated, on the middle panel users parameterize the visual feedback, and on the right panel users chose among various types of sound feedback such as sound replay, signal-based or physically-based sound synthesis.

4.2. Towards Virtual Instrumental Gesture Composition: Assembling and Simulating Units

In this part, we describe experiments that have been conducted for both physics-based synthesis of new percussion performances and sound synthesis. This includes the study of the simulation variability, as well as the assembling and articulation of percussion gesture units for creating new sequences.

4.2.1. Variability in Simulated Movements

An issue that needs to be taken into account is a possible variability of the resultant movements produced by the virtual percussionist.

Depending on the sequence and speed of the gestures chosen, as well as the fine tuning of the mechanical articulations (c.f. equation 3), simulation artifacts can appear and result in large variations in beat impact position and force. These variations are due to the adaptation of the physics simulation to the constraints in the movement data, and can yield unexpected sound phenomena.

An example is presented in Figure 6, where the virtual musician performs a sequence of four beats played legato with different sets of physical constraints (mechanical articulations). On the top graph, a simulation artifact can be seen on the third beat, that results in a larger sound waveform amplitude and sound intensity, actually masking the last beat. This artifact is removed by changing the physics constraints (damping and stiffness coefficients), as shown on the bottom graph. One can see that the four beats are performed similarly, both in terms of amplitude and intensity[1].

[1] Sounds in both simulations have been obtained using Modalys.

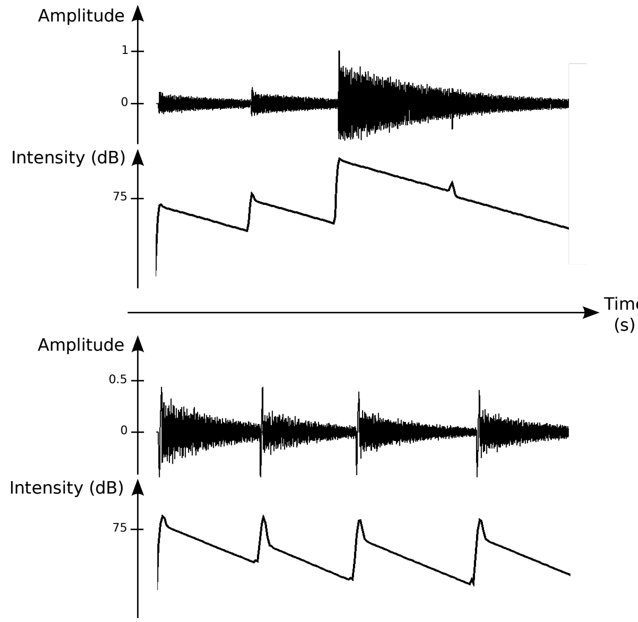


Figure 6. Simulation of four beats played legato. A simulation artifact (top) causes much larger sound waveform amplitude and sound intensity on the third beat, compared to a stable simulation (bottom) generating fairly constant amplitude and intensity for all beats.

4.2.2. Assembly of Gesture Units and Synthesis of new Percussion Sequences

One of the most interesting outcomes of such framework is the possibility of handling and assembling heterogeneous performances using a combination of a few percussion gesture units. Thanks to the physics simulation of the virtual performer, the issue of gesture articulation between movement units is addressed in part by the physics engine, allowing for a more natural sequence of movements[2].

We have tested the simulation of articulating the same percussion gesture (legato) under an accelerando-decelerando musical variation. As shown in Figure 7, the resulting consistency of the instrumental gesture under such a constraint shows the interest of physics simulation. This approach also has an impact on the required size of the movement database obtained through motion capture by reducing the number of different movement articulations that need to be captured.

The current database allows for a combination of the five gesture unit variations described above, where each can be

[2] The resulting articulation is not necessarily equivalent to real performer techniques. It will nevertheless be a physically plausible solution to the problem.

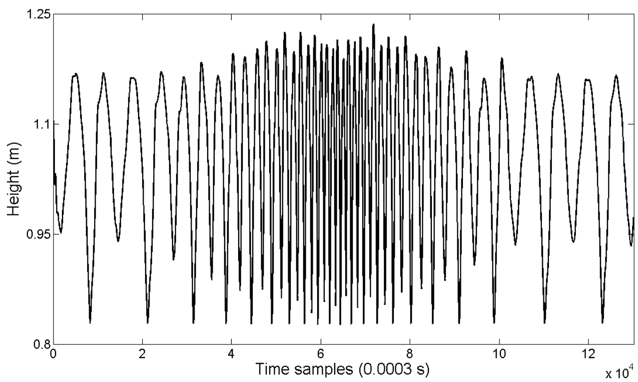
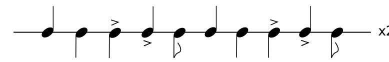


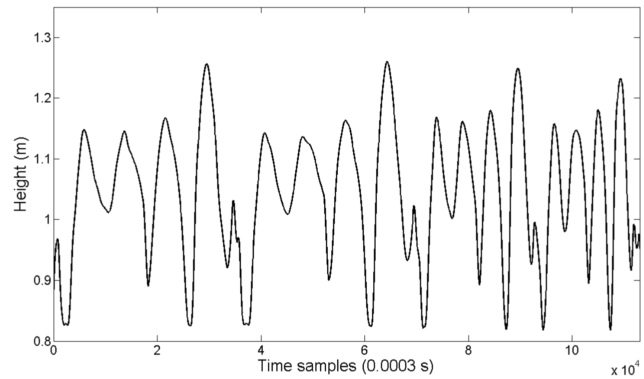
Figure 7. Simulation and articulation between legato beats under an accelerando-decelerando musical variation (height of the tip of the drumstick during the simulation).

performed at different positions on the membrane (centre, one-third and rim) and at various tempi.

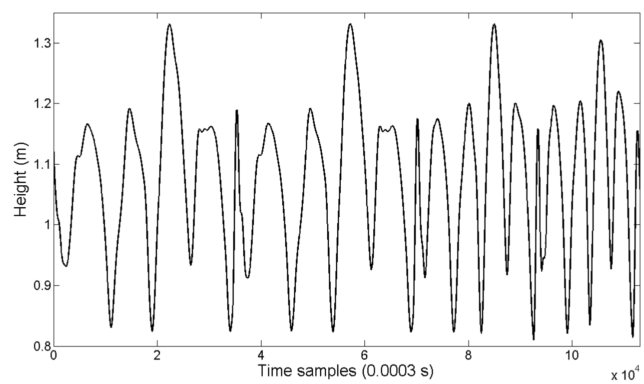
A second simulation example shows the performance of the score presented on the top of Figure 8. Here, the virtual percussionist is given a task that involves variations in percussion gestures units (legato and accent), as well as in the performance tempo. The bottom part of Figure 8 shows the resulting tip trajectories of the drumsticks during the simulation for each hand. One can notice the variability in the amplitude of gestures (height of the tip of the drumstick) reflecting the different gesture units used, and their durations for the notes performed.



(a) Score



(b) Right drumstick



(c) Left drumstick

Figure 8. Score (a) performed by the virtual percussionist, and the resulting height of the right (b) and left (c) tip trajectories of the drumsticks during the simulation.

5. DISCUSSION AND CONCLUSION

5.1. Advantages and Limitations

Simulating instrumental percussion gestures for controlling sound synthesis processes presents advantages and limitations. These are of different orders, with respect to the combination of motion capture data and simulation, the mapping between simulation and sound synthesis, and finally the creation of new percussion performances.

Associating a motion capture database and the physics-based synthesis of instrumental gestures yields an accurate simulation of movements and the retrieval of dynamic aspects of instrumental gestures (beat impacts for percussion performance). It also allows to go beyond pre-recorded motion data by exploring and assembling pre-recorded performance conditions. However, the drawback of such an approach lies in the possible addition of unexpected artifacts to the realism of the captured motion by the physics simulation due to the hard tuning of mechanical joints and/or convergence issues of the numerical algorithms. For instance, increasing the speed of the simulation can introduce errors in the beat impact position or variability in the performance (c.f. Figure 6). Furthermore, our framework depends on the motion clips initially recorded, so that performer tracks are intrinsically part of the resulting simulations. It can be an advantage when considering that performance styles will be preserved, but a limitation when recorded errors are then propagated in the resulting simulations.

Such a physics-based approach considers the mapping between the simulation of percussion gestures and sound synthesis at the physics level. Apart from allowing for a direct mapping between gesture and sound, this choice has the advantage of making clear how physics-based sound synthesis processes can be excited. Moreover, not only physics features can be extracted from the simulation, but also kinematic and energy features. The latter can be of interest for more expressive and complex mappings.

Finally, the synthesis of new percussion performances from the assembly of gesture units yields an intuitive and high level way of controlling a responsive and adaptative virtual percussionist. Nevertheless, assembling instrumental gesture units will not necessarily produce the same level of realism and musical expressiveness as in real performances. This is due to higher-level considerations of instrumental performance that are not taken into account in this model, such as cognitive and expressive strategies, or co-articulation between gesture units. Furthermore, synthesizing a large spectrum of gesture variations necessitates the capture of a large database and the development of adapted motion editing techniques to assembly motion units in a realistic way. Such a physics-based approach however partially eliminates the time-consuming step of building an exhaustive motion capture database, since the simulation itself provides an accurate approximation (in a physics simulation sense) of the gesture co-articulation phenomenon.

5.2. Conclusion and Future Work

We presented in this paper a framework for synthesizing virtual percussion performances by using a motion capture database and a physically-enabled environment in which a virtual percussionist can interact with sound synthesis processes. Our approach is shown to be effective for creating adaptive and responsive virtual percussion conditions and to provide physically relevant inputs to sound synthesis models using physical models of drums.

Future work includes the study of the effect of timpani playing variations on the presented parameters. We aim namely at analyzing the effect of beat locations and musical variations on timpani instrumental gestures. We also plan to enrich the current database with examples of gesture dynamics (*pp*, *mf* and *ff*), as well as with data from different performers.

6. ACKNOWLEDGMENTS

The authors would like to thank F. Marandola (McGill University) and S. Dahl (Aalborg University) for their expert insights on percussion performance, as well as R. Caussé and N. Ellis (IRCAM) for support on Modalys.

This work is partially supported by the *Natural Sciences and Engineering Research Council of Canada (Discovery and Special Research Opportunity grants)* and the *International Doctoral College* of the *Université Européenne de Bretagne*.

7. REFERENCES

[1] J. M. Adrien, “Etude de Structures Complexes Vibrantes, Applications à la Synthèse par Modèles Physiques,” Ph.D. dissertation, Université de Paris VI, France, 1988.

[2] R. Aimi, “Hybrid Percussion: Extending Physical Instruments using Sampled Acoustics,” Ph.D. dissertation, Massachusetts Institute of Technology, USA, 2006.

[3] F. Avanzini and D. Rochesso, “Controlling Material Properties in Physical Models and Sounding Objects,” in *Proc. of the International Computer Music Conference*, 2001, pp. 91–94.

[4] ——, “Physical Modeling of Impacts: Theory and Experiments on Contact Time and Spectral Centroid,” in *Proc. of the International Conference on Sound and Music Computing*, 2004, pp. 287–293.

[5] R. Boie, M. Mathews, and A. Schloss, “The Radio Drum as a Synthesizer Controller,” in *Proc. of the International Computer Music Conference*, 1989, pp. 42–45.

[6] A. Bouënard, S. Gibet, and M. M. Wanderley, “Hybrid Motion Control conbining Inverse Kinematics and Inverse Dynamics Controllers for Simulating Percussion Gestures,” to appear in *Proceedings of the International Conference on Computer Animation and Social Agents*, 2009.

[7] ——, “Real-Time Simulation and Interaction of Percussion Gestures with Sound Synthesis,” technical report, *HAL Open Archives*, 2009.

[8] ——, “Enhancing the Visualization of Percussion Gestures by Virtual Character Animation,” in *Proc. of the International Conference on New Instruments for Musical Expression*, 2008, pp. 38–43.

[9] A. Bouënard, M. M. Wanderley, and S. Gibet, “Analysis of Percussion Grip for Physically Based Character Animation,” in *Proc. of the International Conference on Enactive Interfaces*, 2008, pp. 22–27.

[10] C. Cadoz and C. Ramstein, “Capture, Representation and Composition of the Instrumental Gesture,” in *Proc. of the International Computer Music Conference*, 1990, pp. 53–56.

[11] K. Chuchacz, S. O’Modhrain, and R. Woods, “Physical Models and Musical Controllers: Designing a Novel Electronic Percussion Instrument,” in *Proc. of the International Conference on New Instruments for Musical Expression*, 2007, pp. 37–40.

[12] P. R. Cook, *Real Sound Synthesis For Interactive Applications*. A K Peters Ltd., 2002.

[13] S. Dahl, “On the beat: Human Movement and Timing in the Production and Perception of Music,” Ph.D. dissertation, KTH Royal Institute of Technology, Sweden, 2005.

[14] C. Dobrian and D. Koppelman, "The 'E' in NIME: Musical Expression with New Computer Interfaces," in *Proc. of the International Conference on New Instruments for Musical Expression*, 2006, pp. 277–282.

[15] N. Ellis, J. Bensoam, and R. Caussé, "Modalys Demonstration," in *Proc. of the International Computer Music Conference*, 2005, pp. 101–102.

[16] S. Gibet, "Codage, Représentation et Traitement du Geste Instrumental," Ph.D. dissertation, Institut National Polytechnique de Grenoble, France, 1987.

[17] S. Gibet and P. F. Marteau, "Gestural Control of Sound Synthesis," in *Proc. of the International Computer Music Conference*, 1990, pp. 387–391.

[18] R. Hänninen, L. Savioja, and T. Takala, "Virtual Concert Performance - Synthetic Animated Musicians playing in an Acoustically Simulated Room," in *Proc. of the International Computer Music Conference*, 1996, pp. 402–404.

[19] A. Kapur, P. Essl, G.and Davidson, and P. Cook, "The Electronic Tabla Controller," *Journal of New Music Research*, vol. 32, no. 4, pp. 351–360, 2003.

[20] W. Lytle, "Pipe Dream," *SIGGRAPH 2001 Electronic Theatre*, SIGGRAPH Computer Animation Festival, 2001.

[21] E. R. Miranda and M. M. Wanderley, *New Digital Musical Instruments: Control and Interaction Beyond the Keyboard*. A-R Editions., 2006.

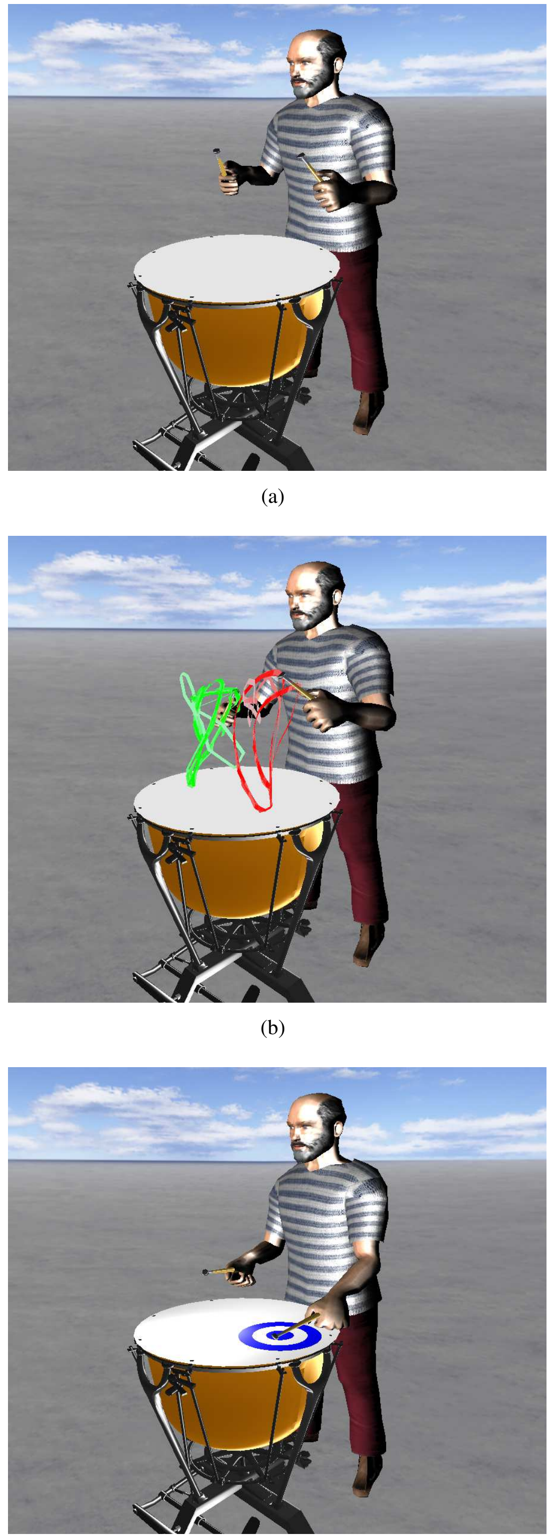
[22] J. O'Brien, C. Shen, and C. Gatchalian, "Synthesizing Sounds from Rigid Body Simulations," in *Proc. of the ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, 2002, pp. 175–181.

[23] G. Rule, "Keyboard Reports: Korg Wavedrum," *Keyboard*, vol. 21, no. 3, pp. 72–78, 1995.

[24] R. Smith, "Open Dynamics Engine," *http://www.ode.org*, 2009.

[25] A. Tindale, A. Kapur, G. Tzanetakis, P. Driessen, and A. Schloss, "A Comparison of Sensor Strategies for Capturing Percussive Gestures," in *Proc. of the International Conference on New Instruments for Musical Expression*, 2005, pp. 200–203.

[26] K. van den Doel, P. Kry, and D. Pai, "FoleyAutomatic: Physically-Based Sound Effects for Interactive Simulation and Animation," in *Proc. of the ACM SIGGRAPH Conference on Computer Graphics and Interactive Techniques*, 2001, pp. 537–544.



(a)

(b)

(c)

Figure 9. Visual feedback during the simulation: users can explore and visualize the virtual percussion performance space (a), as well as drumstick trajectories (b) and beat impact locations (c).